Negligible Genetic Diversity of *Mycobacterium tuberculosis* Host Immune System Protein Targets: Evidence of Limited Selective Pressure

James M. Musser,*,† Amol Amin† and Srinivas Ramaswamy†

**Laboratory of Human Bacterial Pathogenesis, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, Montana 59840 and †Institute for the Study of Human Bacterial Pathogenesis, Department of Pathology, Baylor College of Medicine, Houston, Texas 77030*

Manuscript received September 15, 1999
Accepted for publication January 12, 2000

ABSTRACT

A common theme in medical microbiology is that the amount of amino acid sequence variation in proteins that are targets of the host immune system greatly exceeds that found in metabolic enzymes or other housekeeping proteins. Twenty-four *Mycobacterium tuberculosis* genes coding for targets of the host immune system were sequenced in 16 strains representing the breadth of genomic diversity in the species. Of the 24 genes, 19 were invariant and only six polymorphic nucleotide sites were identified in the 5 genes that did have variation. The results document the highly unusual circumstance that prominent *M. tuberculosis* antigenic proteins have negligible structural variation worldwide. The data are best explained by a combination of three factors: (i) evolutionarily recent global dissemination in humans, (ii) lengthy intracellular quiescence, and (iii) active replication in relatively few fully immunocompetent hosts. The very low level of amino acid diversity in antigenic proteins may be cause for optimism in the difficult fight to control global tuberculosis.

IT is conventional wisdom in medical microbiology that pathogen extracellular and surface proteins involved in interaction with the host immune system and other variable environmental factors are highly polymorphic relative to metabolic enzymes or other "housekeeping" proteins (Selander *et al.* 1994; Li *et al.* 1995; Mathiesen *et al.* 1997; Yamaguchi and Gojobori 1997; Rich *et al.* 1998; Stockbauer *et al.* 1998). Extensive study of polymorphisms in diverse viral, prokaryotic, and eukaryotic pathogens, including those affecting plants and invertebrate and vertebrate animals, have shown this to be the case. The increased variability of surface and extracellular proteins is usually attributed to antigenic diversity to escape host immune functions.

Sequence analysis of 26 structural genes in 842 strains (2 Mb) of the important human pathogen *Mycobacterium tuberculosis* identified a striking reduction of silent nucleotide substitutions and unselected amino acid replacements compared with other microbes (Sreevatsan *et al.* 1997). The lack of variation in structural genes indicated that *M. tuberculosis* is evolutionarily young and has recently spread globally, perhaps as recently as 20,000 years ago (Kapur *et al.* 1994). Although many of the genes encoded metabolic enzymes, the analysis included genes for a 16-kD antigen, a 65-kD heat-shock protein, and catalase-peroxidase, three proteins that are targets of the host immune response (Andersen 1994; Andersen and Brennan 1994; Lyashchenko *et al.* 1998). The paucity of naturally occurring amino acid polymorphisms in these antigens suggested that variation in proteins that interact with the host immune system was uncommon. Inasmuch as a general lack of variation in *M. tuberculosis* antigens would be unprecedented and have important implications for vaccine, therapeutics, and evolutionary biology research, we investigated this issue in detail.

This article is dedicated to Professor R. K. Selander on the occasion of his retirement.

Corresponding author: James M. Musser, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 903 South 4th St., Hamilton, MT 59840.
E-mail: jmusser@niaid.nih.gov

MATERIALS AND METHODS

Bacterial strains: All *M. tuberculosis* isolates used in this study were obtained from the culture collection of J. M. Musser (Sreevatsan *et al.* 1997). A core group of 16 strains (Table 1) was used for initial sequencing of all target genes (Table 2). Although 12 of the 16 organisms were isolated from ethnically diverse patients in Texas, all three principal genetic groups of *M. tuberculosis* were represented. The three principal genetic groups were identified on the basis of the combinations of polymorphisms located in codon 463 of the *katG* gene encoding catalase-peroxidase and codon 95 of *gyrA* encoding the A subunit of DNA gyrase (Sreevatsan *et al.* 1997). In addition, the 16 strains have from 4 to 20 copies of IS*6110* and a diverse array of spoligotypes (Figure 1; Groenen *et al.* 1993; van Embden *et al.* 1993). Diversity in spoligotype and IS*6110* copy number and location in the chromosome have been used extensively as indicators of overall genomic differentiation in *M. tuberculosis* (Groenen *et al.* 1993; van Embden *et al.* 1993). On the basis of molecular analysis of >5000 strains from global sources, the sample represents the breadth of genomic diversity in *M. tuberculosis* strict sense, although it may not fully represent all minor branches of the phylogenetic tree. The

TABLE 1

Characteristics of *M. tuberculosis* isolates analyzed

Strain[a]	Principal genetic group[b]	IS*6110* designation[c]	No. of IS*6110* bands	Spoligotype designation[c]
HN1543	1	003	20	S1
HN1489	1	007	10	S1
HN1343	1	015	9	S1
HN1140	1	033	18	S1
NHN656	1	W44	19	S1
NHN657	1	W148	18	S1
HN1525	2	004	6	S3
HN1426	2	006	4	S25
HN1250	2	016	9	S214
HN1305	2	999	13	S36
NHN390	2	U	10	S29
HN1322	3	001	12	S24
HN1339	3	067	15	S89
HN1423	3	085	5	S100
HN1290	3	120	10	S79
NHN263	3	S	13	S80

[a] HN isolates were cultures from patients in Houston. However, isolates with these same IS*6110* designations are distributed worldwide and have been recovered from patients on many continents (Sreevatsan *et al.* 1997; see Soini *et al.* 2000). NHN656 is from Kenya, NHN657 is from the former Soviet Union, NHN390 is from Peru, and NHN263 is from Mexico.

[b] Principal genetic group designation was determined on the basis of polymorphisms in codons *katG*463 and *gyrA*95 (Sreevatsan *et al.* 1997).

[c] Arbitrary designation.

core sample of 16 strains was supplemented when necessary with additional *M. tuberculosis* isolates recovered from global sources and maintained in the culture collection of J. M. Musser.

Selection of genes encoding antigens and culture supernatant proteins: Extensive study of the host cellular and humoral immune response and characterization of culture supernatants of organisms grown *in vitro* has identified a relatively large number of proteins (Andersen 1994; Andersen and Brennan 1994; Lyashchenko *et al.* 1998). Although not all genes encoding antigenic or supernatant proteins were sequenced in this study, those chosen for analysis represent a random sample of proteins described in the *M. tuberculosis* literature.

Selection of open reading frames (ORFs) encoding members of the Pro-Glu (PE) and Pro-Pro-Glu (PPE) families: The H37Rv genome (Cole *et al.* 1998) contains two large families of genes encoding proteins with PE ($n = 99$) and PPE ($n = 68$) motifs. Four genes of each class were selected for sequence analysis on the basis of being located in positions arrayed around the *M. tuberculosis* chromosome. In addition, one member of the PPE gene family (Rv*3135*) was selected for analysis because alignment of the sequences available for H37Rv and CSU93 (as of June 1999) suggested the gene was variable in size among *M. tuberculosis* isolates.

DNA sequence analysis and PCR size variation in the *Rv3135* gene: Automated DNA sequencing methods using genomic DNA samples and an Applied Biosystems (Foster City, CA) model 377 instrument have been described previously (Sreevatsan *et al.* 1997). The sequence data were assembled and edited using EDITSEQ and MEGALIGN programs (DNASTAR, Madison, WI) and the sequences were compared with published data (Cole *et al.* 1998).

Size variation in the Rv*3135* gene was studied by PCR with the following primers: forward, 5′-TCGACTGCCATACAAC CTG-3′ and reverse, 5′-GTGCTGGTCGAGAACTGAATG-3′, located 210 bp upstream from the Rv*3135* start site and 23 bp downstream from the stop codon, respectively. These primers amplify a 632-bp product from strain H37Rv.

RESULTS

Invariant antigen genes: Of the 24 genes encoding proteins known to be targets of the host immune system, 19 were invariant in the core group of 16 *M. tuberculosis* isolates sequenced. The 19 invariant genes encoded the following proteins: 10-kD culture filtrate protein (CFP10) (Berthet *et al.* 1998); ESAT-6 antigen (Sørensen *et al.* 1995; Ravn *et al.* 1999); 14-kD antigen (Matthews *et al.* 1985; Engers *et al.* 1986; Jackett *et al.* 1988; Verbon *et al.* 1992); MPT63 (16-kD antigen) (Horwitz *et al.* 1995; Lee and Horwitz 1995, 1999); MPT83 (Hewinson *et al.* 1996; Wiker *et al.* 1998); 38-kD lipoprotein antigen (Andersen and Hansen 1989; Espitia *et al.* 1992; Harboe and Wiker 1992); 8.4-kD antigen (Coler *et al.* 1998); CFP17, CFP21, CFP22, CFP25, and CFP29 (Rosenkrands *et al.* 1998; Weldingh *et al.* 1998); MTC28 (28-kD antigen) (Manca *et al.* 1997); superoxide dismutase (23-kD antigen) (Young *et al.* 1985; Zhang *et al.* 1991; Kang *et al.* 1998); antigen 85A (Havlir *et al.* 1991; Wiker and Harboe 1992; Horwitz *et al.* 1995; Garbe *et al.* 1996; Huygen *et al.* 1996; Belisle *et al.* 1997; Lozes *et al.* 1997; Jackson *et al.* 1999); MPT70 antigen (Matsumoto *et al.* 1995); 35-kD antigen (Baird *et al.* 1988; Billman-Jacobe *et al.* 1990); GroES (10-kD antigen) (Cohen *et al.* 1987); and heparin-binding hemagglutinin (Menozzi *et al.* 1996, 1998; Delogu and Brennan 1999).

Variable antigen genes: The remaining five antigen-encoding genes (discussed in detail below) had nucleotide polymorphisms in the 16 *M. tuberculosis* core isolates. One of the nucleotide changes was silent (synonymous substitution), and six would result in amino acid replacements (nonsynonymous substitutions).

45/47-kD secreted antigen complex (MPT32): The 45/47-kD secreted antigen complex is present in culture filtrates of virulent *M. tuberculosis* isolates (Laqueyrerie *et al.* 1995). The antigen complex was initially purified and characterized because of its ability to react with antibodies present in the sera of guinea pigs immunized with live *M. tuberculosis* (Laqueyrerie *et al.* 1995). Diagbouga *et al.* (1997) reported that 40% of sera obtained from smear-positive tuberculosis patients in Burkina Faso reacted with the 45/47-kD antigen complex. We identified one nonsynonymous substitution that resulted in a Phe136Leu amino acid replacement in all principal genetic group 1 isolates and virtually all group 2 organisms. All genetic group 3 *M. tuberculosis* had Phe136.

TABLE 2

Regions of *M. tuberculosis* genes analyzed for nucleotide sequence variation

Rv designation	Size (bp)	Product	Reference	Nucleotides sequenced, positions	No. of sites with synonymous substitutions	No. of sites with nonsynonymous substitutions
Rv0009	546	CFP22 (probable peptidyl-prolyl *cis-trans* isomerase)	Weldingh *et al.* (1998)	−87 to 672	0	0
Rv0040c	930	MTC28 (28-kD proline-rich antigen)	Manca *et al.* (1997)	−137 to 1001	0	0
Rv0129c	1020	Antigen 85C (fibronectin-binding protein C)	Wiker and Harboe (1992)	−137 to 1123	0	1
Rv0475	597	Heparin-binding hemagglutinin	Menozzi *et al.* (1996, 1998)	−330 to 767	0	0
Rv0798c	795	CFP29 (culture filtrate protein)	Rosenkrands *et al.* (1998)	−128 to 917	0	0
Rv0934	1122	38-kD lipoprotein antigen (PhoS homologue)	Andersen and Hansen (1989); Espitia *et al.* (1992); Harboe and Wiker (1992)	−152 to 1185	0	0
Rv1174c	330	8.4-kD antigen	Coler *et al.* (1998)	−191 to 556	0	0
Rv1827	486	CFP17 (culture filtrate protein)	Weldingh *et al.* (1998)	−250 to 592	0	0
Rv1860	975	MPT32 (45/47-kD antigen)	Laqueyrerie *et al.* (1995); Bardou *et al.* (1996); Diagbouga *et al.* (1997)	−183 to 1064	0	1
Rv1886c	975	Antigen 85B (fibronectin-binding protein B)	Wiker and Harboe (1992)	−250 to 999	1	0
Rv1926c	684	MPT63 (16-kD antigen)	Horwitz *et al.* (1995); Lee and Horwitz (1995, 1999)	−104 to 556	0	0
Rv1980c	477	MPT64 (23.5-kD antigen, secreted immunogenic protein)	Harboe *et al.* (1986); Oettinger and Andersen (1994); Oettinger *et al.* (1995); Elhay *et al.* (1998)	−138 to 750	0	1
Rv1984c	651	CFP21 (culture filtrate protein, probable cutinase precursor)	Weldingh *et al.* (1998)	−124 to 865	0	0
Rv2031c	432	14-kD antigen (α-crystallin-like heatshock protein)	Engers *et al.* (1986); Jackett *et al.* (1988); Matthews *et al.* (1985); Verbon *et al.* (1992)	−76 to 545	0	0
Rv2744c	810	35-kD antigen	Baird *et al.* (1988); Billman-Jacobe *et al.* (1990)	−96 to 911	0	0
Rv2873	660	MPT83 (25-kD lipoprotein antigen)	Hewinson *et al.* (1996); Wiker *et al.* (1998)	−72 to 905	0	0
Rv2875	579	MPT70 (major secreted immunogenic protein)	Matsumoto *et al.* (1995)	−107 to 693	0	0
Rv3418c	300	10-kD antigen (GroES homologue)	Cohen *et al.* (1987)	−318 to 373	0	0
Rv3452	678	CFP25 culture filtrate protein (probable cutinase precursor)	Weldingh *et al.* (1998)	−131 to 776	0	0
Rv3763	477	19-kD antigen (lipoprotein antigenic protein)	Harris *et al.* (1994); Lathigra *et al.* (1996); Erb *et al.* (1998)	−251 to 630	0	0
Rv3804c	1014	Antigen 85A (fibronectin-binding protein A)	Wiker and Harboe (1992)	−235 to 1035	0	0
Rv3846	621	23-kD antigen (superoxide dismutase)	Young *et al.* (1985); Zhang *et al.* (1991); Kang *et al.* (1998)	−105 to 674	0	0

(*Continued*)

TABLE 2
(Continued)

Rv designation	Size (bp)	Product	Reference	Nucleotides sequenced, positions	No. of sites with synonymous substitutions	No. of sites with nonsynonymous substitutions
Rv3874	300	CFP-10 (culture filtrate protein)	Berthet *et al.* (1998)	−142 to 358	0	0
Rv3875	285	ESAT-6 (early secretory antigenic protein)	Sørensen *et al.* (1995); Ravn *et al.* (1999)	−266 to 396	0	0
Rv0285	306	PE family protein	Cole *et al.* (1998)	−120 to 378	0	0
Rv1169	300	PE family protein	Cole *et al.* (1998)	−93 to 355	0	0
Rv1788	297	PE family protein	Cole *et al.* (1998)	−84 to 369	0	0
Rv3477	294	PE family protein	Cole *et al.* (1998)	−185 to 511	0	0
Rv0388c	540	PPE family protein	Cole *et al.* (1998)	−232 to 658	0	0
Rv1790	1050	PPE family protein	Cole *et al.* (1998)	−100 to 1133	0	1
Rv2430	582	PPE family protein	Cole *et al.* (1998)	−140 to 692	0	0
Rv3135	396	PPE family protein	Cole *et al.* (1998)	−230 to 463	Variable	0

19-kD antigen: The 19-kD antigen is a highly expressed surface-associated glycolipoprotein that is a dominant antigen in infected humans (Andersen and Brennan 1994; Harris *et al.* 1994; Lathigra *et al.* 1996; Erb *et al.* 1998; Mahenthiralingam *et al.* 1998). Its role in human-pathogen interactions is not known. Lathigra *et al.* (1996) suggested that the protein is a virulence factor, but inactivation of the homologous *Mycobacterium intracellulare* gene failed to demonstrate an essential role in growth and virulence in BALB/c mice (Mahenthiralingam *et al.* 1998). Our analysis identified two amino acid replacements, Gly115Arg and Cys158Ser, both occurring in a distinct phylogenetic lineage of principal genetic group 2 organisms. The Cys158Ser replacement was found in strain HN1489; both the Gly115Arg and Cys158Ser amino acid changes were identified in strain HN1426. The Cys158Ser amino acid residue is located in a region of the molecule containing a mouse linear B-cell epitope (Harris *et al.* 1994). However, neither polymorphism is located in a region known to be recognized by human sera or in T-cell epitopes (Lathigra *et al.* 1996; Erb *et al.* 1998). Sequence analysis of the 19-kD antigen gene in 47 additional principal genetic group 2 *M. tuberculosis* strains confirmed that these amino acid variants were confined to clonally related organisms having similar IS*6110* profiles and four or six hybridizing copies of this insertion element.

MPT64 (23.5-kD antigen): MPT64 is a 23.5-kD secreted antigen that is identical to a protein initially purified from culture filtrates of *Mycobacterium bovis* BCG Tokyo and designated MPB64 (Harboe *et al.* 1986; Oettinger and Andersen 1994). Humans mount an immune response to MPT64 during tuberculosis disease. T-cell and linear B-cell epitope mapping has been conducted with recombinant proteins (Oettinger *et al.* 1995). A putative delayed-type hypersensitivity-inducing epitope is located within amino acid residues Gly-173 to Ala-187 (Oettinger *et al.* 1995; Elhay *et al.* 1998). Our analysis identified a Thr35Ala amino acid replacement in one strain from Peru (NHN390) that is a member of principal genetic group 2. We note that Oettinger and Andersen (1994) reported that monoclonal antibody C24b3 defined an epitope composed of two structural domains located in the sequences Ala-1 to Leu-43 and Ala-108 to Ser-152.

Antigen-85B and antigen-85C proteins: The antigen-85 complex is formed by three major proteins—designated Ag85A, Ag85B, and Ag85C—that are expressed by actively replicating organisms (Horwitz *et al.* 1995). Ag85-complex proteins are mycolyltransferases that participate in the final stages of cell wall synthesis (Wiker and Harboe 1992; Belisle *et al.* 1997; Jackson *et al.* 1999). Their expression is upregulated in response to isoniazid treatment *in vitro* (Harris *et al.* 1994; Bardou *et al.* 1996). Ag85-complex proteins induce cellular and humoral immune responses in infected laboratory animals and humans (Wiker and Harboe 1992). Levels of anti-Ag85 antibodies are usually low in healthy purified

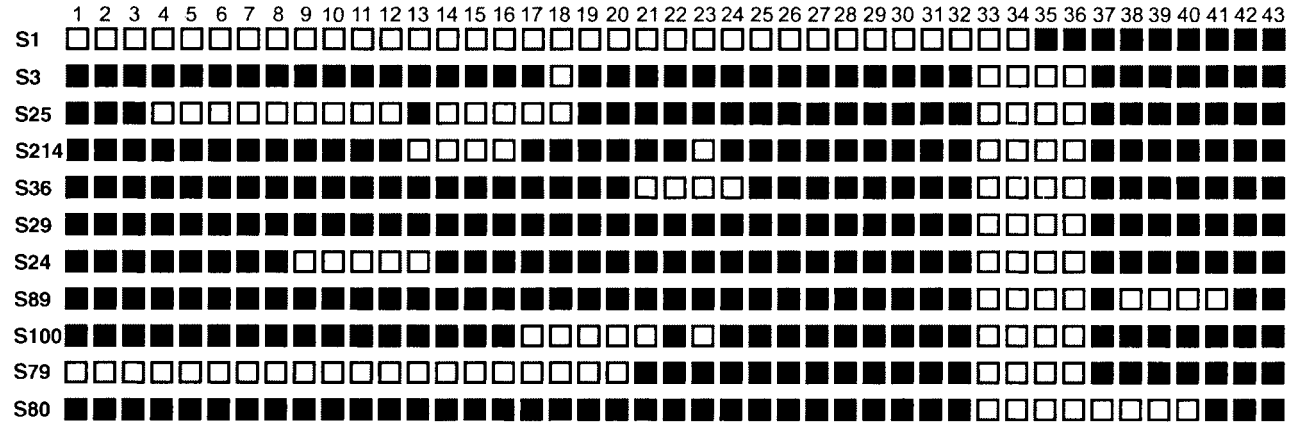

Figure 1.—Spoligotype patterns identified among the 16 *M. tuberculosis* isolates studied. The arbitrary spoligotype pattern designations (Soini *et al.* 2000) are shown at left. Solid squares, hybridization with the designated spacer probe; open squares, lack of hybridization.

protein-derivative (PPD) positive individuals but increase in patients with active tuberculosis (Havlir *et al.* 1991; Wiker and Harboe 1992; Lyashchenko *et al.* 1998). Active immunization with purified Ag85-complex proteins or DNA encoding Ag85B can induce protective immunity and a positive therapeutic effect in laboratory animals (Huygen *et al.* 1996; Lozes *et al.* 1997; Lowrie *et al.* 1999). The precise role of these proteins in *M. tuberculosis* host-pathogen interactions is not yet clear, but they are known to be upregulated in response to growth in macrophages *in vitro* (Lee and Horwitz 1995). Recently, Lee and Horwitz (1999) mapped T-cell epitopes in outbred guinea pigs immunized with purified Ag85A and Ag85B.

One synonymous (silent) nucleotide change was found in codon 238 (CCC → CCA; Pro ↔ Pro) of the Ag85B gene in HN1543. In addition, all five organisms belonging to principal genetic group 3 had a G → A nucleotide substitution located at position −3 relative to the start codon. Sequence analysis of 20 additional isolates found that all 14 members of principal genetic group 3, regardless of IS*6110* pattern or spoligotype, also had this upstream polymorphism. In contrast, this nucleotide substitution was not present in members of the other two principal genetic groups. All 16 core isolates had the same sequence for the gene encoding Ag85C except strain HN1305, which had a G → A mutation at position −63 upstream of the start codon and a nucleotide change resulting in a Glu103Asp amino acid replacement.

PE and PPE genes: The *M. tuberculosis* H37Rv genome contains two large families of genes encoding glycine-rich proteins designated PE (Pro-Glu) and PPE (Pro-Pro-Glu) (Cole *et al.* 1998). The H37Rv genome contains 99 members of the PE family and 68 members of the PPE family. The cell location and biologic function of these proteins are unknown, but it has been suggested that they participate in antigenic variation or interfere with host immune responses by inhibiting antigen processing (Cole *et al.* 1998). Recently, Dillon *et al.* (1999) reported that a 39-kD member of the PPE family elicited strong T-cell proliferative and gamma-interferon responses in peripheral blood mononuclear cells from PPD-positive individuals. No nucleotide variation was identified in five (Rv*0285*, Rv*0388*, Rv*1169*, Rv*1788*, and Rv*2430*) of the eight PE and PPE genes chosen for sequence analysis. Two isolates belonging to principal genetic group 1 (HN1343 and HN1489) had a C → T nucleotide polymorphism located at position −78 relative to the start codon of Rv*3477*. One group 2 (HN1305) and one group 3 (HN1339) organism had a Pro190Ser amino acid replacement in Rv1790.

Rv3135, a member of the PPE family of proteins, was uniquely variable. PCR analysis of the Rv*3135* gene in the core group of 16 isolates identified four distinct sizes of products, and sequence analysis revealed them to be 322 bp, 632 bp, 1021 bp, and 1973 bp long, respectively (Figure 2). Isolates assigned to principal genetic group 1 had the 1973-bp or 1021-bp sequences, those belonging to group 2 had the 632-bp or 322-bp sequences, and those belonging to group 3 had the 632-bp variant. To more fully investigate the extent and phylogenetic distribution of Rv*3135* size variation, PCR analysis was also conducted on 141 *M. tuberculosis* isolates from global sources. Four additional PCR fragment sizes were found, and sequence analysis identified products of 200 bp, 460 bp, 499 bp, and 533 bp, respectively (Figure 2). As observed for the 16 core isolates, all group 3 organisms had the 632-bp product and group 2 organisms had either the 322-bp or 632-bp sequence. In contrast, group 1 organisms had a broad range of Rv*3135* gene sizes, including 200 bp, 460 bp, 499 bp, 533 bp, 1021 bp, and 1973 bp. Inspection of all Rv*3135* data found that, in principle, each variant could be linked to one or more of the other variants by a single molecular step, suggesting that the Rv*3135* polymorphisms were generated by rare deletion or insertion events. On the basis of the inferred amino acid sequences, not all iso-

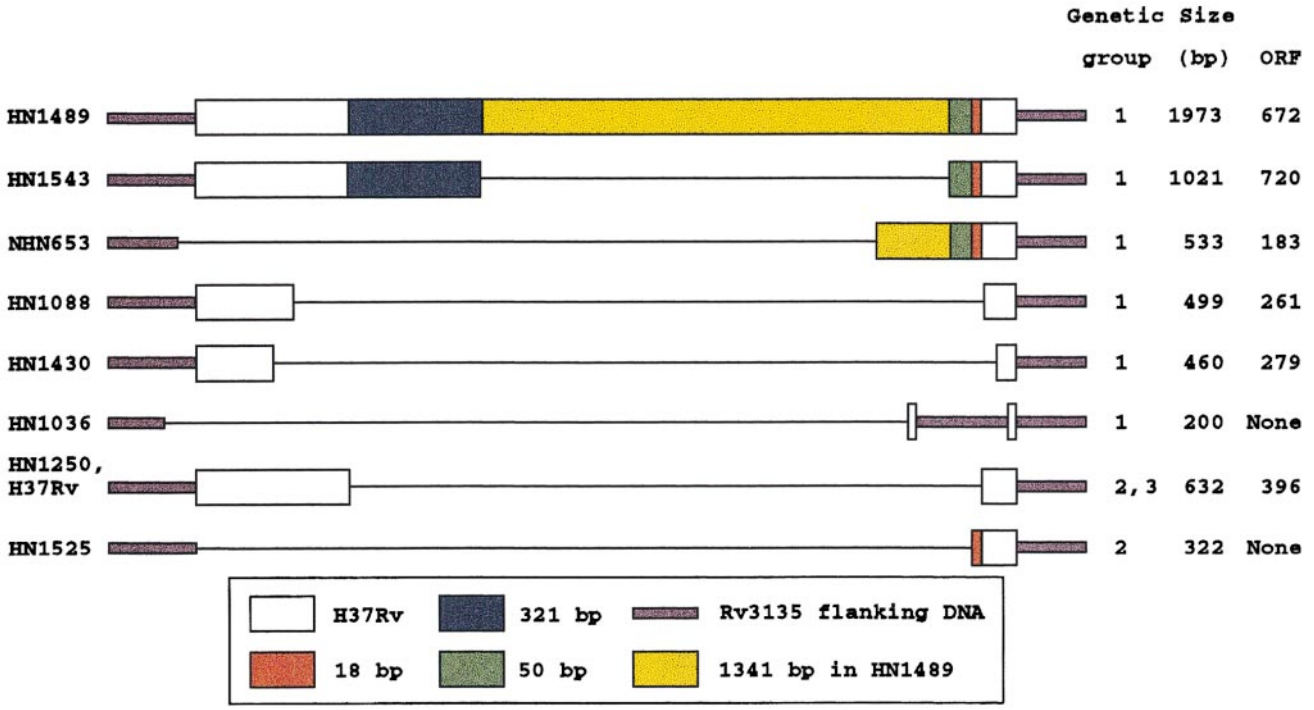

Figure 2.—Schematic representation of variation identified in the Rv*3135* gene encoding a member of the Pro-Pro-Glu (PPE) family of *M. tuberculosis* proteins. The thin horizontal line represents DNA found in the 1973-bp segment of HN1489 but absent from other strains. The genetic group was defined by polymorphisms present in *katG* codon 463 and *gyrA* codon 95 (Sreevatsan *et al.* 1997). The Rv*3135* gene was obtained by PCR from representative isolates. The sizes (in base pairs) shown are the lengths of the PCR products obtained for each variant. The lengths of the longest open reading frame for each variant are shown at right. Strains HN1250, HN1489, HN1525, and HN1543 are described in Table 1. Strains HN1036, HN1088, and HN1430, recovered from patients in Houston, Texas, are principal genetic group 1 (Sreevatsan *et al.* 1997) and have spoligotype S1 (Soini *et al.* 2000).

lates would express a protein product (Figure 3). Some of the gene variants had premature stop codons resulting in truncated proteins relative to the H37Rv product. In addition, two of the variants lacked upstream regions encoding putative regulatory sequences and, hence, would probably not express the Rv3135 protein.

DISCUSSION

Most of the 26 *M. tuberculosis* structural genes previously studied (Sreevatsan *et al.* 1997) for allelic variation encode intracellular metabolic enzymes that are not known to directly interact with the host. However, allelic variation was also restricted in three genes encoding proteins (16-kD antigen, 65-kD heat-shock protein, and catalase-peroxidase) that are host immune system targets (Andersen 1994; Andersen and Brennan 1994; Lyashchenko *et al.* 1998). Inasmuch as an important common theme in infectious disease research is that pathogen surface and other proteins involved in direct interaction with the host immune system are highly polymorphic relative to metabolic housekeeping enzymes (Selander *et al.* 1994; Li *et al.* 1995; Mathiesen *et al.* 1997; Yamaguchi and Gojobori 1997; Rich *et al.* 1998; Stockbauer *et al.* 1998), we thought it critical to more fully investigate the level of antigen gene variation. The analysis was also motivated by the substantial interest in protein-subunit vaccine strategies and other new therapeutic and diagnostic methods (Andersen 1994; Horwitz *et al.* 1995; Huygen *et al.* 1996; Tascon *et al.* 1996; Lozes *et al.* 1997; Harth and Horwitz 1999). In addition, since many of the proteins released into the extracellular medium from growing *M. tuberculosis* are immunologically recognized during infection in animals and early stages of human pulmonary tuberculosis (Pal and Horwitz 1992; Orme *et al.* 1993; Andersen 1994; Andersen and Brennan 1994; Horwitz *et al.* 1995), the failure of tuberculosis patients to uniformly respond to many of those same proteins raised the possibility that amino acid sequence variation in the target molecules is a contributing factor.

Our analysis demonstrated that an absence or very low frequency of structural polymorphism is a general characteristic of genes encoding prominent targets for host B-cell and T-cell immune responses. There are several hypotheses that may account for this observation: (i) the diversification of antigens is restricted by the intracellular niche, (ii) the organism has a low spon-

```
HN1489   MDYAFLPPEINSARMYSGPGPNSMLVAAASWDALAAELAS  40
HN1543   MDYAFLPPEINSARMYSGPGPNSMLVAAASWDALAAELAS  40
H37Rv    MDYAFLPPEINSARMYSGPGPNSMLVAAASWDALAAELAS  40
HN1430   MDYAFLPPEINSARMYSGPGPNSMLVAAASWDALAAELAS  40
HN1088   MDYAFLPPEINSARMYSGPGPNSMLVAAASWDALAAELAS  40

HN1489   AAENYGSVIARLTGMHWWGPASTSMLAMSAPYVEWLERTA  80
HN1543   AAENYGSVIARLTGMHWWGPASTSMLAMSAPYVEWLERTA  80
H37Rv    AAENYGSVIARLTGMHWWGPASTSMLAMSAPYVEWLERTA  80
HN1430   AAENYGSVIARLTGMHWWGPASKRPVPATPLALHLDPFSS  80
HN1088   AAENYGSVIARLTGMHWWGPASTSMLAIRGAIVVETASAS  80

HN1489   AQTKQTATQARAAAAAFEQAHAMTVPPALVTANRAELKAL  120
HN1543   AQTKQTATQARAAAAAFEQAHAMTVPPALVTANRAELKAL  120
H37Rv    AQTKQTATQARAAAAAFEQAHAMTVPPALVTGIRGAIVVE  120
HN1430   RPARHRIRTNVRS                             120
HN1088   NTAGTPP                                   120

HN1489   IASNLLGQNTAAIAAHRGTVRRDVGTRRGRDVRLRHHLSG  160
HN1543   IASNLLGQNTAAIAAIEAQYAEMWAQDAAAMYGYATTSAA  160
H37Rv    TASASNTAGTPP                              160

HN1489   GETVDAVLLATTDHQPGRASRPERRGHPGRHQLRREHADR  200
HN1543   ARQLTPFSSPQQTTNPAGLAAQNAAVTQAATNSAGNTPTA  200

HN1489   IVATVLFPVAGSRGADGMAQHSPG                  240
HN1543   LSQLSSFLSQAVEAPTGWRSTATSGHMTWITTVLTRQPDR  240


NHN653   MPFADAWIPSPIASAKRSCHHTARHGSANPRHTPQLRPGH  40
NHN653   DTGPRHASGRTTRVQLPRAAI                     80
```

Figure 3.—Alignment of amino acid residues in Rv3135 variants. Aligned are the inferred amino acid sequences for each of the six Rv3135 protein variants putatively expressed by *M. tuberculosis* isolates and identified in this study. Two of the Rv*3135* gene PCR variants (200 bp and 322 bp, see Figure 2) may not result in expression of a protein product due to structural alterations (apparent deletions) of the presumed upstream regulatory sequence. The inferred amino acid sequence for NHN653 is shown at the bottom of the figure because the amino terminus does not align with the other proteins. Strains HN1489 and HN1543 are described in Table 1. Strains HN1088 and HN1430, recovered from patients in Houston, Texas, are principal genetic group 1 (Sreevatsan *et al.* 1997) and have spoligotype S1 (Soini *et al.* 2000).

taneous mutation rate, (iii) humans have little effective immunity to the target proteins so far characterized, (iv) major protective immune targets have not yet been identified, and (v) the widespread dissemination of *M. tuberculosis* is evolutionarily very recent.

In principle the intracellular lifestyle of *M. tuberculosis* may restrict antigen variation by sequestering the pathogen from the host immune response, especially the humoral defense mechanism. However, there is no evidence that restricted antigen gene variation is a common theme for intracellular pathogens, including obligate parasites in the extreme case (Weidmann *et al.* 1997; Brisse *et al.* 1998; Stothard *et al.* 1998). For example, abundant levels of nucleotide and amino acid polymorphisms have been described in *Chlamydia trachomatis* (Stothard *et al.* 1998), *Salmonella enterica* (Selander *et al.* 1994; Li *et al.* 1995), *Listeria monocytogenes* (Weidmann *et al.* 1997), and *Trypanosoma cruzi* (Brisse *et al.* 1998). In addition, evidence indicates that humoral immune processes participate in host defense against *M. tuberculosis* (Glatman-Freedman and Casadevall 1998; Teitelbaum *et al.* 1998). Hypothesis ii is ruled out because *M. tuberculosis* does not have an unusually low spontaneous mutation frequency (David and Newman 1971). On the contrary, spontaneous drug resistance mutants can be selected rapidly in the laboratory and in human patients (David and Newman 1971; Bifani *et al.* 1996; Ramaswamy and Musser 1998). In regard to hypothesis iii, immunity to *M. tuberculosis* is complex, involving innate factors (Bellamy *et al.* 1998) and acquired humoral and cellular responses. Without documented *in vitro* or animal model correlates of human protective immunity, it is difficult to speculate about the potential contributions of hypotheses iii and iv, above. However, generation of a protective or therapeutic response to tuberculosis by immunization with several of the antigens we studied has been well described in animal models, which reduces the likelihood that these two hypotheses alone are responsible. Evidence has also been presented that patients with pulmonary tuberculosis have depressed effector immune function, which can be long lasting (Hirsch *et al.* 1996, 1999a,b). Effective immunity could also be altered by the magnitude of posttranslational modification of target proteins by, for example, glycosylation. According to this hypothesis, glycosylation would alter amino acid residues encompassed in a region of the molecule commonly recognized as foreign. This modification would then alter host immune system recognition of the target protein. In this regard, it was recently shown that deglycosylation of the 45/47-kD antigen complex significantly decreased its capacity to elicit *in vivo* and *in vitro* cellular immune responses (Horn *et al.* 1999; Romain *et al.* 1999). We also note that Herrmann *et al.* (1996) postulated that glycosylation functions to regulate the relative amounts of cell-associated and soluble forms of protein antigens. In addition, effective immunity may depend on response to nonprotein antigens, and Teitelbaum *et al.* (1998) reported that a monoclonal antibody specific for arabinomannan conferred partial protection on mice challenged by the aerosol route. Finally, certain *M. tuberculosis* isolates may not be rapidly recognized as foreign early in the course of host-pathogen interactions (Manca *et al.* 1999).

It is highly unlikely that an immunologic selection hypothesis alone accounts for our observations. One strong argument favoring hypothesis v (the very recent origin and widespread dissemination of *M. tuberculosis*) is the absence or very low level of silent nucleotide substitutions in *all* genes characterized to date, which in our laboratory now number ~80 genes characterized in large numbers of strains from global sources. Moreover, comparison of the genome sequences available for H37Rv and CSU93 indicate that although several large insertion and deletion events differentiate the genomes, none of the roughly 4000 ORFs is strongly polymorphic at the nucleotide level. The lack of significant variation in genes encoding proteins that are host immune system targets indicates that in all likelihood *M. tuberculosis* has spread globally even more recently than roughly 20,000 years ago (Kapur *et al.* 1994; Sreevatsan

et al. 1997). If, as the data indicate, interaction of humans with *M. tuberculosis* has been frequent for only a brief evolutionary period, then immunogenetics may play an even more prominent role in determining the outcome of the host-pathogen relationship. The occurrence of epidemic waves of tuberculosis in many populations in recent centuries and in certain inbred and historically isolated human populations (Sousa *et al.* 1997) is consistent with this idea.

Early in the course of host-pathogen interaction *M. tuberculosis* is sequestered in the macrophage. Most individuals control the organism without detrimental effects. The pathogen then resides in a quiescent state with relatively little replication until cell-mediated immunity is compromised, commonly by HIV infection or aging. Most other pathogens are frequently confronted with very labile environments in which there is a substantial premium on variation and adaptation. In contrast, the life cycle of *M. tuberculosis* lessens the extent of exposure to wide fluctuations in variable host and environmental factors and minimizes DNA replication cycles. Together, these two factors tend to constrain species diversity. Importantly, emergence from the quiescent state usually occurs when the host is relatively immunocompromised, which means that the pathogen usually is not exposed to strong immune selective pressure. Indeed, it well may be that this aspect of the *M. tuberculosis* infection cycle is a primary cause of the lack of variation in the protein antigens studied, in spite of being responsible for 3 million deaths yearly. These considerations, coupled with the evolutionarily recent introduction of *M. tuberculosis* into humans from nonhuman hosts living in close association, such as cattle, goats, or water buffalo, may contribute to the lack of amino acid polymorphism in this pathogen.

PE and PPE variation: Cole *et al.* (1998) speculated that the PE and PPE protein families are a source of antigenic variation in *M. tuberculosis.* Of note, we found that the gene encoding PPE protein Rv3135 varied in size, although not all of the allelic variants would result in a protein product. Size variants were not found in high frequency in isolates recovered from the same patient over time, suggesting that the gene is not hypervariable or hypermutable (S. Ramaswamy and J. M. Musser, unpublished data). The nonrandom association of Rv3135 variants with distinct genetic groups also is consistent with a lack of a hypervariable or hypermutable phenotype.

There are 167 PE and PPE gene family members in the H37Rv chromosome (Cole *et al.* 1998). Variation was identified in the coding sequence of 1 of the 8 PE and PPE genes analyzed, suggesting by extrapolation that roughly 20 PE and PPE members would be polymorphic in *M. tuberculosis.* Alignment of additional members of the PE and PPE families in H37Rv and CSU93 indicates that other examples of size polymorphism exist. Notwithstanding the important recent observation that strains may lack certain PE or PPE genes (Gordon *et al.* 1999), the results suggest that a relatively small percentage of these proteins are variable. However, this issue requires more extensive investigation at the gene and protein level in large samples of isolates.

Implications for development of new therapeutics and diagnostics: The current global tuberculosis epidemic and the spread of multidrug-resistant strains make development of new therapeutics an important priority. Recently, Harth and Horwitz (1999) demonstrated that an inhibitor of exported *M. tuberculosis* glutamine synthetase selectively blocked pathogen growth in axenic culture and human monocytes. Their findings suggest that extracellular proteins made by *M. tuberculosis* are potential novel drug targets; a similar idea was advanced for the Ag85-complex mycolyltransferase (Belisle *et al.* 1997). Our findings, together with data for the 26 genes previously reported (Sreevatsan *et al.* 1997), indicate that prominent *M. tuberculosis* antigenic proteins and potential novel drug and diagnostic targets are likely to have very little structural variation worldwide. If the apparent lack of host immune pressure can be abrogated by induction of enhanced effector function, then the lack of diversity in protein targets may be cause for optimism in the difficult fight to control global tuberculosis.

We thank colleagues who generously contributed strains. We are indebted to A. Casadevall, M. L. Gennaro, B. N. Kreiswirth, L. Schlesinger, and R. K. Selander for reading the manuscript and suggesting improvements. This research was supported by U.S. Public Health Services grants AI-41168, AI-37004, and DA-09238 to J.M.M.

LITERATURE CITED

Andersen, P., 1994 Effective vaccination of mice against *Mycobacterium tuberculosis* infection with a soluble mixture of secreted mycobacterial proteins. Infect. Immun. **62:** 2536–2544.

Andersen, Å. B., and P. Brennan, 1994 Proteins and antigens of *Mycobacterium tuberculosis*, pp. 307–332 in *Tuberculosis: Pathogenesis, Protection, and Control*, edited by B. R. Bloom. American Society for Microbiology, Washington, DC.

Andersen, Å. B., and E. B. Hansen, 1989 Structure and mapping of antigenic domains of protein antigen b, a 38,000-molecular-weight protein of *Mycobacterium tuberculosis.* Infect. Immun. **57:** 2481–2488.

Baird, P. N., L. M. C. Hall and A. R. M. Coates, 1988 A major antigen from *Mycobacterium tuberculosis* which is homologous to the heat shock proteins groES from *E. coli* and the htpA gene product of *Coxiella burneti.* Nucleic Acids Res. **16:** 9047.

Bardou, F., A. Quémard, M.-A. Dupont, C. Horn, G. Marchal *et al.*, 1996 Effects of isoniazid on ultrastructure of *Mycobacterium aurum* and *Mycobacterium tuberculosis* and on production of secreted proteins. Antimicrob. Agents Chemother. **40:** 2459–2467.

Belisle, J. T., V. D. Vissa, T. Sievert, K. Takayama, P. J. Brennan *et al.*, 1997 Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. Science **276:** 1420–1422.

Bellamy, R., C. Ruwende, T. Corrah, K. P. McAdam, H. C. Whittle *et al.*,, 1998 Variations in the *NRAMP1* gene and susceptibility to tuberculosis in West Africans. N. Engl. J. Med. **338:** 640–644.

Berthet, F.-X., P. B. Rasmussen, I. Rosenkrands, P. Andersen and B. A. Gicquel, 1998 *Mycobacterium tuberculosis* operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10). Microbiol. **144:** 3195–3203.

Bifani, P., B. B. Plikaytis, V. Kapur, K. Stockbauer, X. Pan *et al.*,

1996 Origin and interstate spread of a New York City multidrug-resistant *Mycobacterium tuberculosis* clone family. J. Am. Med. Assoc. **275:** 452–457.

Billman-Jacobe, H., A. J. Radford, J. S. Rothel and P. R. Wood, 1990 Mapping of the T and B cell epitopes of the *Mycobacterium bovis* protein, MPB70. Immunol. Cell. Biol. **68:** 359–365.

Brisse, S., C. Barnabé, A. L. Bañuls, I. Sidibé, S. Noël *et al.*, 1998 A phylogenetic analysis of the *Trypanosoma cruzi* genome project CL Brener reference strain by multilocus enzyme electrophoresis and multiprimer random amplified polymorphic DNA fingerprinting. Mol. Biochem. Parasitol. **92:** 253–263.

Cohen, M. L., L. W. Mayer, H. S. Rumschlag, M. A. Yakrus, W. D. Jones, Jr., *et al.*, 1987 Expression of proteins of *Mycobacterium tuberculosis* in *Escherichia coli* and potential of recombinant genes and proteins for development of diagnostic reagents. J. Clin. Microbiol. **25:** 1176–1180.

Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher *et al.*, 1998 Deciphering the biology of *Mycobacterium tuberculosis* from the complete genomc sequence. Nature **393:** 537–544.

Coler, R. N., Y. A. W. Skeiky, T. Vedvick, T. Bement, P. Ovendale *et al.*, 1998 Molecular cloning and immunologic reactivity of a novel low molecular mass antigen of *Mycobacterium tuberculosis.* J. Immunol. **161:** 2356–2364.

David, H. L., and C. M. Newman, 1971 Some observations on the genetics of isoniazid resistance in the tubercle bacilli. Am. Rev. Respir. Dis. **104:** 508–515.

Delogu, G., and M. J. Brennan, 1999 Functional domains present in the mycobacterial hemagglutinin, HBHA. J. Bacteriol. **181:** 7464–7469.

Diagbouga, S., F. Fumoux, A. Zoubga, P. T. Sanou and G. Marchal, 1997 Immunoblot analysis for serodiagnosis of tuberculosis using a 45/47-kilodalton antigen complex of *Mycobacterium tuberculosis.* Clin. Diag. Lab. Immunol. **4:** 334–338.

Dillon, D. C., M. R. Alderson, C. H. Day, D. M. Lewinsohn, R. Coler *et al.*, 1999 Molecular characterization and human T-cell responses to a member of a novel *Mycobacterium tuberculosis mtb39* gene family. Infect. Immun. **67:** 2941–2950.

Elhay, M. J., T. Oettinger and P. Andersen, 1998 Delayed-type hypersensitivity responses to ESAT-6 and MPT64 from *Mycobacterium tuberculosis* in the guinea pig. Infect. Immun. **66:** 3454–3456.

Engers, H. D., V. Houba, J. Bennedsen, T. M. Buchanan, S. D. Chaparas *et al.*, 1986 Results of a World Health Organization-sponsored workshop to characterize antigens recognized by mycobacterium-specific monoclonal antibodies. Infect. Immun. **51:** 718–720.

Erb, K. J., J. Kirman, L. Woodfield, T. Wilson, D. M. Collins *et al.*, 1998 Identification of potential $CD8^+$ T-cell epitopes of the 19 kD and AhpC proteins of *Mycobacterium tuberculosis.* No evidence for $CD8^+$ T-cell priming against the identified peptides after DNA-vaccination of mice. Vaccine **16:** 692–697.

Espitia, C., M. Elinos, R. Hernández-Pando and R. Mancilla, 1992 Phosphate starvation enhances expression of the immunodominant 38-kilodalton protein antigen of *Mycobacterium tuberculosis*: demonstration by immunogold electron microscopy. Infect. Immun. **60:** 2998–3001.

Garbe, T. R., N. S. Hibler and V. Deretic, 1996 Isoniazid induces expression of the antigen 85 complex in *Mycobacterium tuberculosis.* Antimicrob. Agents Chemother. **40:** 1754–1756.

Glatman-Freedman, A., and A. Casadevall, 1998 Serum therapy for tuberculosis revisited: reappraisal of the role of antibody-mediated immunity against *Mycobacterium tuberculosis.* Clin. Microbiol. Rev. **11:** 514–532.

Gordon, S. V., R. Brosch, A. Billault, T. Garnier, K. Eiglmeier *et al.*, 1999 Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. Mol. Microbiol. **32:** 643–655.

Groenen, P. M. A., A. E. Bunschoten, D. van Soolingen and J. D. A. van Embden, 1993 Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*: application for strain differentiation by a novel typing method. Mol. Microbiol. **10:** 1057–1065.

Harboe, M., and H. G. Wiker, 1992 The 38-kD protein of *Mycobacterium tuberculosis*: a review. J. Infect. Dis. **166:** 874–884.

Harboe, M., S. Nagai, M. E. Patarroyo, M. L. Torres, C. Ramirez *et al.*, 1986 Properties of proteins MPB64, MPB70, and MPB80 of *Mycobacterium bovis* BCG. Infect. Immun. **52:** 293–302.

Harris, D. P., H.-M. Vordermeier, S. J. Brett, G. Pasvol, C. Moreno *et al.*, 1994 Epitope specificity and isoforms of the mycobacterial 19-kilodalton antigen. Infect. Immun. **62:** 2963–2972.

Harth, G., and M. A. Horwitz, 1999 An inhibitor of exported *Mycobacterium tuberculosis* glutamine synthetase selectively blocks the growth of pathogenic mycobacteria in axenic culture and in human monocytes: extracellular proteins as potential novel drug targets. J. Exp. Med. **189:** 1425–1435.

Havlir, D. V., R. S. Wallis, W. H. Boom, T. M. Daniel and K. Chervenak, 1991 Human immune response to *Mycobacterium tuberculosis* antigens. Infect. Immun. **59:** 665–670.

Herrmann, J. L., P. O'Gaora, A. Gallagher, J. E. R. Thole and D. B. Young, 1996 Bacterial glycoproteins: a link between glycosylation and proteolytic cleavage of a 19 kD antigen from *Mycobacterium tuberculosis.* EMBO J. **15:** 3547–3554.

Hewinson, R. G., S. L. Michell, W. P. Russell, R. A. McAdam and W. R. Jacobs, Jr., 1996 Molecular characterization of MPT83: a seroreactive antigen of *Mycobacterium tuberculosis* with homology to MPT70. Scand. J. Immunol. **43:** 490–499.

Hirsch, C. S., R. Hussain, Z. Toossi, G. Dawood, F. Shahid *et al.*, 1996 Cross-modulation by transforming growth factor β in human tuberculosis: suppression of antigen-driven blastogenesis and interferon γ production. Proc. Natl. Acad. Sci. USA **93:** 3193–3198.

Hirsch, C. S., Z. Toossi, C. Othieno, J. L. Johnson, S. K. Schwander *et al.*, 1999a Depressed T-cell interferon-γ responses in pulmonary tuberculosis: analysis of underlying mechanisms and modulation with therapy. J. Infect. Dis. **180:** 2069–2073.

Hirsch, C. S., Z. Toossi, G. Vanham, J. L. Johnson, P. Peters *et al.*, 1999b Apoptosis and T cell hyporesponsiveness in pulmonary tuberculosis. J. Infect. Dis. **179:** 945–953.

Horn, C., A. Namane, P. Pescher, M. Riviere, F. Romain *et al.*, 1999 Decreased capacity of recombinant 45/47-kD molecules (Apa) of *Mycobacterium tuberculosis* to stimulate T lymphocyte responses related to changes in their mannosylation pattern. J. Biol. Chem. **274:** 32023–32030.

Horwitz, M. A., B.-W. E. Lee, B. J. Dillon and G. Harth, 1995 Protective immunity against tuberculosis induced by vaccination with major extracellular proteins of *Mycobacterium tuberculosis.* Proc. Natl. Acad. Sci. USA **92:** 1530–1534.

Huygen, K., J. Content, O. Denis, D. L. Montgomery, A. M. Yawman *et al.*, 1996 Immunogenicity and protective efficacy of a tuberculosis DNA vaccine. Nat. Med. **2:** 893–898.

Jackett, P. S., G. H. Bothamley, H. V. Batra, A. Mistry, D. B. Young *et al.*, 1988 Specificity of antibodies to immunodominant mycobacterial antigens in pulmonary tuberculosis. J. Clin. Microbiol. **26:** 2313–2318.

Jackson, M., C. Raynaud, M.-A. Lanéelle, C. Guilhot, C. Laurent-Winter *et al.*, 1999 Inactivation of the antigen 85C gene profoundly affects the mycolate content and alters the permeability of the *Mycobacterium tuberculosis* cell envelope. Mol. Microbiol. **31:** 1573–1587.

Kang, S.-K., Y.-J. Jung, C.-H. Kim and C.-Y. Song, 1998 Extracellular and cytosolic iron superoxide dismutase from *Mycobacterium* bovis BCG. Clin. Diag. Lab. Immunol. **5:** 784–789.

Kapur, V., T. S. Whittam and J. M. Musser, 1994 Is *Mycobacterium tuberculosis* 15,000 years old? J. Infect. Dis. **170:** 1348–1349.

Laqueyrerie, A., P. Militzer, F. Romain, K. Eiglmeier, S. Cole *et al.*, 1995 Cloning, sequencing, and expression of the *apa* gene coding for the *Mycobacterium tuberculosis* 45/47-kilodalton secreted antigen complex. Infect. Immun. **63:** 4003–4010.

Lathigra, R., Y. Zhang, M. Hill, M.-J. Garcia, P. S. Jackett *et al.*, 1996 Lack of production of the 19-kD glycolipoprotein in certain strains of *Mycobacterium tuberculosis.* Res. Microbiol. **147:** 237–249.

Lee, B.-Y., and M. A. Horwitz, 1995 Identification of macrophage and stress-induced proteins of *Mycobacterium tuberculosis.* J. Clin. Invest. **96:** 245–249.

Lee, B.-Y., and M. A. Horwitz, 1999 T-cell epitope mapping of the three most abundant extracellular proteins of *Mycobacterium tuberculosis* in outbred guinea pigs. Infect. Immun. **67:** 2665–2670.

Li, J., H. Ochman, E. A. Groisman, E. F. Boyd, F. Solomon *et al.*, 1995 Relationship between evolutionary rate and cellular location among the Inv/Spa invasion proteins of *Salmonella enterica.* Proc. Natl. Acad. Sci. USA **92:** 7252–7256.

Lowrie, D. B., R. E. Tascon, V. L. D. Bonato, V. M. F. Lima, L. H.

Faccioli *et al.*, 1999 Therapy of tuberculosis in mice by DNA vaccination. Nature **400:** 269–271.

Lozes, E., K. Huygen, J. Content, O. Denis, D. L. Montgomery *et al.*, 1997 Immunogenicity and efficacy of a tuberculosis DNA vaccine encoding the components of the secreted antigen 85 complex. Vaccine **15:** 830–833.

Lyashchenko, K., R. Colangeli, M. Houde, H. Al Jahdali, D. Menzies *et al.*, 1998 Heterogeneous antibody responses in tuberculosis. Infect. Immun. **66:** 3936–3940.

Mahenthiralingam, E., B.-I. Marklund, L. A. Brooks, D. A. Smith, G. Bancroft *et al.*, 1998 Site-directed mutagenesis of the 19-kilodalton lipoprotein antigen reveals no essential role for the protein in the growth and virulence of *Mycobacterium tuberculosis*. Infect. Immun. **66:** 3626–3634.

Manca, C., K. Lyashchenko, R. Colangeli and M. L. Gennaro, 1997 MTC28, a novel 28-kilodalton proline-rich secreted antigen specific for the *Mycobacterium tuberculosis* complex. Infect. Immun. **65:** 4951–4957.

Manca, C., L. Tsenova, C. E. Barry III, A. Bergtold, S. Freeman *et al.*, 1999 *Mycobacterium tuberculosis* CDC1551 induces a more vigorous host response in vivo and in vitro, but is not more virulent than other clinical isolates. J. Immunol. **162:** 6740–6746.

Mathiesen, D. A., J. H. Oliver, Jr., C. P. Kolbert, E. D. Tullson, B. J. Johnson *et al.*, 1997 Genetic heterogeneity of *Borrelia burgdorferi* in the United States. J. Infect. Dis. **175:** 98–107.

Matsumoto, S., T. Matsuo, N. Ohara, H. Hotokezaka, M. Naito *et al.*, 1995 Cloning and sequencing of a unique antigen MPT70 from *Mycobacterium tuberculosis* H37Rv and expression in BCG using *E. coli*-mycobacteria shuttle vector. Scand. J. Immunol. **41:** 281–287.

Matthews, R., A. Scoging and A. D. M. Rees, 1985 Mycobacterial antigen-specifc human T-cell clones secreting macrophage activating factors. Immunology **54:** 17–23.

Menozzi, F. D., J. H. Rouse, M. Alavi, M. Laude-Sharp, J. Muller *et al.*, 1996 Identification of a heparin-binding hemagglutinin present in mycobacteria. J. Exp. Med. **184:** 993–1001.

Menozzi, F. D., R. Bischoff, E. Fort, M. J. Brennan and C. Locht, 1998 Molecular characterization of the mycobacterial heparin-binding hemagglutinin, a mycobacterial adhesin. Proc. Natl. Acad. Sci. USA **95:** 12625–12630.

Oettinger, T., and Å. B. Andersen, 1994 Cloning and B-cell-epitope mapping of MPT64 from *Mycobacterium tuberculosis* H37Rv. Infect. Immun. **62:** 2058–2064.

Oettinger, T., A. Holm, I. M. Mtoni, Å. B. Andersen and K. Hasløv, 1995 Mapping of the delayed-type hypersensitivity-inducing epitope of secreted protein MPT64 from *Mycobacterium tuberculosis*. Infect. Immun. **63:** 4613–4618.

Orme, I. M., P. Andersen and W. H. Boom, 1993 T cell response to *Mycobacterium tuberculosis*. J. Infect. Dis. **167:** 1481–1497.

Pal, P. G., and M. A. Horwitz, 1992 Immunization with extracellular proteins of *Mycobacterium tuberculosis* induces cell-mediated immune responses and substantial protective immunity in a guinea pig model of pulmonary tuberculosis. Infect. Immun. **60:** 4781–4792.

Ramaswamy, S., and J. M. Musser, 1998 Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. Tuberc. Lung Dis. **79:** 3–29.

Ravn, P., A. Demissie, T. Eguale, H. Wondwosson, D. Lein *et al.*, 1999 Human T cell responses to the ESAT-6 antigen from *Mycobacterium tuberculosis*. J. Infect. Dis. **179:** 637–645.

Rich, S. M., M. C. Licht, R. R. Hudson and F. J. Ayala, 1998 Malaria's Eve: evidence of a recent population bottleneck throughout the world populations of *Plasmodium falciparum*. Proc. Natl. Acad. Sci. USA **95:** 4425–4430.

Romain, F., C. Horn, P. Pescher, A. Namane, M. Riviere *et al.*, 1999 Deglycosylation of the 45/47-kilodalton antigen complex of *Mycobacterium tuberculosis* decreases its capacity to elicit in vivo or in vitro cellular immune responses. Infect. Immun. **67:** 5567–5572.

Rosenkrands, I., P. B. Rasmussen, M. Carnio, S. Jacobsen, M. Theisen *et al.*, 1998 Identification and characterization of a 29-kilodalton protein from *Mycobacterium tuberculosis* culture filtrate recognized by mouse memory effector cells. Infect. Immun. **66:** 2728–2735.

Selander, R. K., J. Li, E. F. Boyd, F.-S. Wang and K. Nelson, 1994 DNA sequence analysis of the genetic structure of populations of *Salmonella enterica* and *Escherichia coli*, pp. 17–49 in *Bacterial Diversity and Systematics*, edited by F. G. Priest, A. Ramos-Cormenzana and B. J. Tindall. Plenum, New York.

Soini, H., X. Pan, A. Amin, E. A. Graviss, A. Siddiqui *et al.*, 2000 Characterization of *Mycobacterium tuberculosis* isolates from patients in Houston, Texas, by spoligotyping. J. Clin. Microbiol. **38:** 669–676.

Sørensen, A. L., S. Nagai, G. Houen, P. Andersen and Å. B. Andersen, 1995 Purification and characterization of a low-molecular-mass T-cell antigen secreted by *Mycobacterium tuberculosis*. Infect. Immun. **63:** 1710–1717.

Sousa, A. O., J. I. Salem, F. K. Lee, M. C. Vercosa, P. Cruaud *et al.*, 1997 An epidemic of tuberculosis with a high rate of tuberculin anergy among a population previously unexposed to tuberculosis, the Yanomami Indians of the Brazilian Amazon. Proc. Natl. Acad. Sci. USA **94:** 13227–13232.

Sreevatsan, S., X. Pan, K. E. Stockbauer, N. D. Connell, B. N. Kreiswirth *et al.*, 1997 Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. Proc. Natl. Acad. Sci. USA **94:** 9869–9874.

Stockbauer, K. E., D. Grigsby, X. Pan, Y.-X. Fu, L. M. Perea Mejia *et al.*, 1998 Hypervariability generated by natural selection in an extracellular complement-inhibiting protein of serotype M1 strains of group A *Streptococcus*. Proc. Natl. Acad. Sci. USA **95:** 3128–3133.

Stothard, D. R., G. Boguslawski and R. B. Jones, 1998 Phylogenetic analysis of the *Chlamydia trachomatis* major outer membrane protein and examination of potential pathogenic determinants. Infect. Immun. **66:** 3618–3625.

Tascon, R. E., M. J. Colston, S. Ragno, E. Stavropoulos, D. Gregory *et al.*, 1996 Vaccination against tuberculosis by DNA injection. Nat. Med. **2:** 888–891.

Teitelbaum, R., A. Glatman-Freedman, B. Chen, J. B. Robbins, E. Unanue *et al.*, 1998 A mAb recognizing a surface antigen of *Mycobacterium tuberculosis* enhances host survival. Proc. Natl. Acad. Sci. USA **95:** 15688–15693.

van Embden, J. D. A., M. D. Cave, J. T. Crawford, J. W. Dale, K. D. Eisenach *et al.*, 1993 Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J. Clin. Microbiol. **31:** 406–409.

Verbon, A., R. A. Hartskeerl, A. Schuitema, A. H. J. Kolk, D. B. Young *et al.*, 1992 The 14,000-molecular-weight antigen of *Mycobacterium tuberculosis* is related to the alpha-crystallin family of low-molecular-weight heat shock proteins. J. Bacteriol. **174:** 1352–1359.

Weidmann, M., J. L. Bruce, C. Keating, A. E. Johnson, P. L. McDonough *et al.*, 1997 Ribotypes and virulence gene polymorphisms suggest three distinct *Listeria monocytogenes* lineages with differences in pathogenic potential. Infect. Immun. **65:** 2707–2716.

Weldingh, K., I. Rosenkrands, S. Jacobsen, P. B. Rasmussen, M. J. Elhay *et al.*, 1998 Two-dimensional electrophoresis for analysis of *Mycobacterium tuberculosis* culture filtrate and purification and characterization of six novel proteins. Infect. Immun. **66:** 3492–3500.

Wiker, H. G., and M. Harboe, 1992 The antigen 85 complex: a major secretion product of *Mycobacterium tuberculosis*. Microbiol. Rev. **56:** 648–661.

Wiker, H. G., K. P. Lyashchenko, A. M. Aksoy, K. A. Lightbody, J. M. Pollock *et al.*, 1998 Immunochemical characterization of the MPB70/80 and MPB83 proteins of *Mycobacterium bovis*. Infect. Immun. **66:** 1445–1452.

Yamaguchi, Y., and T. Gojobori, 1997 Evolutionary mechanisms and population dynamics of the third variable envelope region of HIV within single hosts. Proc. Natl. Acad. Sci. USA **94:** 1264–1269.

Young, R. A., B. R. Bloom, C. M. Grosskinsky, J. Ivanyi, D. Thomas *et al.*, 1985 Dissection of *Mycobacterium tuberculosis* antigens using recombinant DNA. Proc. Natl. Acad. Sci. USA **82:** 2583–2587.

Zhang, Y., R. Lathigra, T. Garbe, D. Catty and D. Young, 1991 Genetic analysis of superoxide dismutase, the 23 kilodalton antigen of *Mycobacterium tuberculosis*. Mol. Microbiol. **5:** 381–391.

Communicating editor: Y.-X. Fu